

## Cooperation and cultural transmission in a coordination game

Calabuig, Vicente; Olcina, Gonzalo

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

www.peerproject.eu

### Empfohlene Zitierung / Suggested Citation:

Calabuig, V., & Olcina, G. (2009). Cooperation and cultural transmission in a coordination game. *Journal of Economic Behavior & Organization*, 72(1), 188-201. <https://doi.org/10.1016/j.jebo.2009.05.005>

### Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu> Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**gesis**  
Leibniz-Institut  
für Sozialwissenschaften

### Terms of use:

This document is made available under the "PEER Licence Agreement". For more Information regarding the PEER-project see: <http://www.peerproject.eu> This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der  
  
Leibniz-Gemeinschaft

## Accepted Manuscript

Title: Cooperation and cultural transmission in a coordination game

Authors: Vicente Calabuig, Gonzalo Olcina

PII: S0167-2681(09)00135-8  
DOI: doi:10.1016/j.jebo.2009.05.005  
Reference: JEBO 2394

To appear in: *Journal of Economic Behavior & Organization*

Received date: 6-3-2007  
Revised date: 4-5-2009  
Accepted date: 6-5-2009

Please cite this article as: Calabuig, V., Olcina, G., Cooperation and cultural transmission in a coordination game, *Journal of Economic Behavior and Organization* (2008), doi:10.1016/j.jebo.2009.05.005

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



# COOPERATION AND CULTURAL TRANSMISSION IN A COORDINATION GAME.<sup>1</sup>

Vicente Calabuig<sup>2</sup> and Gonzalo Olcina<sup>3</sup>.

Department of Economic Analysis and  
ERI-CES , University of Valencia, Spain.

December, 2008

<sup>1</sup>We thank two anonymous referees for very helpful comments. This paper has also benefited from comments of participants in the XXIX Spanish Symposium of Economic Analysis in Pamplona (Spain), in the 2nd World Congress of the Game Theory Society in Marseille (France) and in the VI Spanish Meeting on Game Theory and Practice in Elche (Spain). The authors acknowledge financial support from the Spanish Ministry of Education and Science Project SEJ2005-08054/ECON and also from the Fundacion BBVA.

<sup>2</sup>Corresponding author: Edifici Departamental Oriental, Avda. dels Tarongers, s/n 46022 Valencia. Spain Tel: +34963828262; e-mail : Vicente.Calabuig@uv.es

<sup>3</sup>Edifici Departamental Oriental, Avda. dels Tarongers, s/n 46022 Valencia. Spain Tel: +34963828241; e-mail : Gonzalo.Olcina@uv.es

### Abstract

The aim of this paper is to analyze if cooperation can be the product of cultural evolution in a two-stage coordination game, consisting of a production stage followed by a negotiation phase. We present an overlapping generations model with cultural transmission of preferences where the distribution of preferences in the population and the strategies are determined endogenously and simultaneously. There are several groups in the society; some of them play cooperatively and others do not. Socialization takes place inside the group, but there is a positive rate of migration among groups which parents anticipate. Our main result shows that all groups converge to the cooperative equilibrium.

**Keywords:** Cultural Transmission, Coordination Game, Social Preferences, Cooperation, Migration.

**JEL classification:** C78; D64; D63.

## 1 Introduction.

Most of the economic, social and biological research trying to explain the appearance, maintenance and evolution of cooperation in human societies has frequently used the prisoner's dilemma as a simple game that exemplifies this central problem. But researchers have neglected the study of another old social dilemma that from our point of view deserves at least as much attention as the prisoner's dilemma: a simple coordination game, known as the stag hunt. This game formalizes a story told by Rousseau of two hunters who could cooperate by jointly hunting a stag or defect by individually hunting a hare. In contrast to a prisoner's dilemma, where defection is the best response regardless of the other's strategy, in stag-hunt games, defection is the best response to defection, but cooperation is the best response to cooperation. Thus, the stag hunt has two equilibria, one where players cooperate and another where they defect.

The viability of cooperation in the society within and among groups depends on mutual beliefs and rests on trust. This crucial dimension of cooperation is much better captured in the stag hunt game than in the prisoner's dilemma. In this latter game there is a conflict between individual rationality and mutual benefits (efficiency), while in the stag hunt there is a conflict between mutual benefits and personal risk. In other words, it is rational to cooperate but you need to trust mutually to do so.

The stag hunt or coordination game has been analyzed in the literature from an evolutionary approach. Most of this work shows that it is difficult to escape the inefficient but risk dominant equilibrium in an evolutionary setting (see, for instance, Kandori et al. 1993, Young 1993 and Ellison 1993). The experimental evidence also shows that cooperative solutions to the coordination problem are not as easy to obtain as one might think (Cooper, 1999). Groups can get "stuck" at a non cooperative equilibrium (van Huyck et al. 1990).

A common feature in these previous works is that they analyze an "isolated" one-shot coordination game; that is, payoffs are uniquely determined by the players' choices when generating surplus, without depending upon any additional action. But most of the economic interactions

that turn out to have a coordination game structure are actually sequential two-stage games. Players firstly make productive decisions, generating a joint surplus, and subsequently, they bargain over its division among them.

A well-known and important example is the hold up problem on specific investments in an incomplete contract scenario. Some authors (see for instance, Ellingsen and Robles 2002, Trogér 2002 and Dawid and McLeod 2001) have investigated two-stage hold up games in an evolutionary framework. In their models, players distribute the obtained surplus between them employing a symmetric and fixed bargaining procedure (a simultaneous Nash demand game).

However, these papers are concerned with the magnitude of investment. We, on the other hand, are concerned with a qualitative aspect of investment. In particular, agents often must choose the degree to which their investments are relationship specific. Of course, the more relationship specific the agent's investment, the worse is that agent's outside option. That is, choices in the production stage influence the allocation of bargaining power in later stages.

In our model, players choose simultaneously between cooperation and non-cooperation in the production stage and thereafter, in the second stage, surplus is distributed through a negotiation mechanism that depends upon behavior in the first stage. If both players cooperate, they have equal bargaining power. If one player cooperates and the other does not, we assume that the latter has all the bargaining power.

Cooperation in this game means specialization, that is, to choose a relation-specific investment or action, while non-cooperation means non specialization or to choose a general-purpose investment or action. Economic surplus arises and increases with specialization and mutual cooperation. But the distribution of the obtained surplus depends on the relative ex post bargaining strength of the partners which is in turn negatively correlated with specialization. This is the original story behind the stag hunt game: each player has to make a decision between specializing in hunting a stag or hunting hare. However, it is needed bilateral cooperation in order to catch the stag with positive probability, while catching a hare is a matter of solo production.

We think that in real life many bilateral joint ventures fit in this framework (for instance,

marriage, joint research...). Labour division and specialization increase total surplus but weaken the bargaining power of the party who specializes. For instance, firms and workers often invest in job-specific assets and job-related training whose returns are shared through subsequent wage negotiations. If investments are non-contractible, this scenario generally results in a trade-off: highly specific investments yield a larger surplus to be divided between the partners but reduce the ex post bargaining position of the investor, provided his partner has chosen a less specific type of investment. There is a conflict between mutual benefit and the individual risk of getting locked in the relationship with a very weak bargaining position and being exploited by the other party in the negotiation stage.

When our two-stage coordination game is played by individuals with self-regarding preferences and solved backwards, it results in a particular version of the stag hunt game known as assurance or mutualism game. The surplus when one player hunts hare (non-cooperation) while the other hunts stag (cooperation) is greater than the surplus obtained when both hunt hare. However, when the game is played by individuals with other-regarding preferences the structure of the game changes and cooperation becomes the unique equilibrium.

In this paper we want to study if the Pareto dominant (cooperative) equilibrium can be obtained (and if so, under which conditions) as the product of cultural evolution. We work with a dynamic model of purposeful and costly cultural transmission, where the distribution of preferences in the population and the strategies in the two-stage coordination game in the long run are determined endogenously and simultaneously.

More precisely, we present an overlapping generations model where agents live for two periods. In the first period they are children who are educated in certain preferences or values. In the second period, they are adults and are randomly matched with another members of the population to play the two-stage coordination game.

Preferences in the population are heterogeneous. In each period there is a fraction of selfish

players and there is also a fraction of players motivated by reciprocal altruism<sup>1</sup>.

Both the educational effort of their parents and the social environment influence a child's preferences. The distribution of preferences in the population evolves depending on the socialization effort of both types of parents, which is determined by two factors: their expectations about the equilibrium to be played in the next generation and the current distribution of preferences (since oblique transmission is a substitute of vertical transmission).

If the proportion of selfish players in the initial condition of the dynamics is high, then there is still multiplicity of equilibria in the game with heterogeneous preferences. We assume that there are several groups in the society; some of them play cooperatively and others do not. Socialization takes place inside the group, but there is a positive rate of migration among groups which parents anticipate.

Our main result is that, for very general conditions, in the globally stable steady state of the society there will be a mixed distribution of preferences where both selfish and other-regarding preferences are present and more importantly, all groups coordinate in the cooperative equilibrium of the stag hunt game. The non cooperative groups eventually end up playing cooperatively because the preference distribution in the group evolves increasing the proportion of inequity averse individuals.

The driving force of the result is that the presence of a significant fraction of inequity averse individuals in the population acts as a stock of social capital in the society. In other words, it works as a good substitute for complete contracting, reducing personal risk, and this occurs because of their aversion to inequality. In a population with a high proportion of inequity averse individuals, their effect is so strong that cooperation is the unique equilibrium. Inequity averse players are rather generous and fair when they have all the bargaining power, and this makes cooperation the best response of selfish players even against defection. On the other hand, inequity averse players credibly threat to punish opportunistic and greedy behaviour by

---

<sup>1</sup>In particular, we use the concept of inequity averse preferences of Fehr and Schmidt (1999).



rejecting unfair offers in the negotiation. This makes selfish players behave also very generously when they have all bargaining power.

Our paper is related to a cultural transmission literature in dynamic models of preference evolution. Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985), in their seminal work in evolutionary anthropology, were the first to propose models of cultural transmission with exogenous socialization efforts. However we follow the class of cultural transmission models first analyzed by Bisin and Verdier (1998) in which the socialization efforts are endogenous. Moreover, Bisin et al. (2004) analyze the evolution of cooperation in a context where players play a prisoner's dilemma and in which there is also heterogeneity of preferences.

Sandholm (2001) presents a two speed dynamic analysis of the evolution of preferences in a 2x2 game. Gradual changes in preferences are accompanied by immediate behavioral adjustments that maintain equilibrium play. He shows how aggregate behavior can jump discretely in an instant of evolutionary time in coordination games. However, in this work, preferences evolve according to an indirect evolutionary approach, although this approach can be interpreted as a cultural transmission model where preferences that lead to economic success are more likely to be transmitted to subsequent generations.

Kuran and Sandholm (2008) analyse a pure coordination game (in material payoffs) in which agents also care about identity-driven personal ideals, captured by an individual-specific preference parameter. For any given preference parameter profile in a community, there is a unique Nash equilibrium in the random matching game. Cultural evolution is driven by parents' transmission and by self-persuasion. The main goal of this paper is the analysis of cultural integration when there is cross-cultural interaction between communities. In particular, any individual has a positive probability of playing the coordination game and culturally interacting with members from the other community.

Our model shares with these papers that we also work with cultural transmission as the engine of preferences and behavior evolution and that we also assume a two speed adjustment process. But there are several relevant differences. On the one hand, instead of analyzing a

pure coordination game, we study a two stage game that results in a coordination game in material payoffs. The second stage, the sharing or negotiation stage, turns out to be crucial for the working of social preferences. On the other hand, our main interest lies in the equilibrium selection problem. We want to know if the interaction of cultural transmission and migration between culturally non-isolated populations yields convergence to the cooperative equilibrium. But in our model immigrants play the equilibrium played in their new community. Therefore, the possibility of migration only influences the socialization effort of parents, affecting in this way the preferences distribution in the original population. This change in preferences can in turn cause a shift in the equilibrium played.

The paper is organized as follows. Section 2 describes the two-stage incomplete information coordination game. In Section 3 we obtain and characterize the equilibria of the game. Section 4 summarizes the mechanism of cultural transmission of preferences and analyzes the optimal education effort choice of the different types of parents. Section 5 presents our main result on migration among groups and cooperation. Finally, we conclude in Section 6.

## 2 A coordination game with a distribution phase.

Players are randomly matched into pairs in each period to play a two-stage game. We denote these stages as the production and distribution stage, respectively.

In the production stage, each player has to decide, independently and simultaneously, whether to cooperate (C) or not to cooperate (NC) and their pair of actions determines the surplus obtained. If both players cooperate it is jointly obtained the highest total surplus, denoted by  $2B$ . If only one of them cooperates it is obtained a surplus of size  $d$ . And, finally, if both players do not cooperate, then they obtain the lowest surplus denoted by  $2b$ . We assume that  $B > d > 2b > 0$ . Therefore, the efficient outcome of this situation is the one in which both players choose the cooperative action.

In the distribution stage, we suppose that, after observing the realized surplus, players

bargain following an ultimatum game with the following characteristics depending on the pair of first stage actions. If both players cooperate, they have equal bargaining power, that is, they get to be the proposer with equal probability. If one player cooperates and the other does not cooperate, we assume that the latter has all the bargaining power, that is, he will be the proposer in the ultimatum game. And finally, if both players do not cooperate, we assume, for simplicity, that no additional action is taken and each one obtains a payoff of  $b$ .

This game represents a bilateral joint venture. As we mentioned in the introduction, cooperation means specialization, that is, to choose a relation-specific action or investment, while non-cooperation means non specialization or to choose a general-purpose action or investment. Therefore, we are assuming quite realistically that economic surplus increases with specialization and mutual cooperation.

The distribution of the generated surplus depends on the relative ex post bargaining strength of the partners which is in turn negatively correlated with specialization. The reason is very intuitive: when a subject makes a general investment, he has a valuable outside option. Conversely, if he makes a specific investment, he has a small outside option. Therefore, in the former case, the player has a high bargaining power and in the latter case he has a low bargaining power<sup>2</sup>.

Assume, as conventional economics and game theory do, that all players have self-regarding preferences and are risk neutral. Selfish players accept any division of the surplus when they are responders in the ultimatum negotiation game. And they offer nothing to the responder when they are the proposers and have all the bargaining power.

Therefore, if we solve the two-stage game by backward induction, we find that the players are facing the following simultaneous game in the production stage:

---

<sup>2</sup>This result is shown formally in a general bargaining model in Olcina and Penyarrubia(2002).

	C	NC
C	$B, B$	$0, d$
NC	$d, 0$	$b, b$

This coordination game is also known as the assurance or mutualism game (because  $d > 2b$ ). For instance, consider an actual stag hunt game. It is assumed that the number (or the size) of the captured hares is greater whenever there is no competition in hunting hare with the other player. This coordination game has two (subgame perfect) Nash equilibria: one in which both players cooperate in the first stage (C, C) and another in which they do not cooperate (NC, NC). Nothing in the rational behavior of the players prevents choosing one of the two equilibria. A clear answer to this problem of equilibrium selection remains elusive.

### **Social preferences.**

Until recently standard game theory has assumed that all players are self-regarding, in the sense that they are only motivated by their own monetary payoff. However, there are many pieces of experimental data that indicates that a significant fraction of the subjects does not care only about absolute material payoffs but rather relative payoffs. These experiments suggest that fairness and reciprocity motives affect the behaviour of many people, but not all people.

A number of theoretical models have been developed in the literature to obtain reciprocal behaviour. Well-known examples include Fehr and Schmidt (1999) and Bolton and Ockenfelds (2000)'s models of inequity aversion, Charness and Rabin (2002)'s model of quasi-maximin preferences, Rabin (1993) and Dufwenberg and Kirchsteiger (2004)'s models of intention-based reciprocity. These models can lead to different predictions in some particular games, but in our two-stage game they all deliver the same qualitative results. We choose the inequity aversion preferences model of Fehr and Schmidt (1999) for tractability reasons.

We assume that in each period  $t$  there is a proportion,  $p_t$ , of self-interested players in the population and another proportion,  $1 - p_t$ , of players motivated by inequity aversion.

Let  $x = (x_1, x_2)$  denote the vector of monetary payoffs for both players. The utility function of an inequity averse player  $i$  is given by:

$$U_i(x) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}, j \neq i \quad (1)$$

where  $\beta_i \leq \alpha_i$  and  $0 \leq \beta_i < 1$ .

Inequity averse agents are willing to give up some material payoff to move in the direction of more equitable outcomes. The second term in (1) measures the utility loss from disadvantageous inequity, while the third term measures the loss from advantageous inequity. The assumption  $\beta_i \leq \alpha_i$  implies that a player suffers more from inequity that is to his disadvantage, that is, the inequity aversion is asymmetric. We are interested in strongly inequity averse players, those with the parameters  $\beta > 0.5$ .<sup>3</sup>

The utility function of selfish players are characterized by the above utility functions with  $\alpha_i = \beta_i = 0$ , that is,  $U_i(x) = x_i$ .

### 3 Cooperation and Equilibrium Behaviour.

In this section we characterize the perfect bayesian equilibria of the two stage incomplete information game described in the previous section. In this game, no player knows the true type of player that he is randomly matched with but he knows the distribution of preferences of the population from which a player is drawn. We will start analyzing the behaviour of both types of players in the distribution (negotiation) phase.

#### 3.1 Behaviour of inequity averse players.

Strongly inequity averse players behave very differently in a negotiation as compared to selfish players. In particular they are very generous as proposers and they are willing to reject greedy

---

<sup>3</sup>We also assume that the following condition holds for the inequity averse players:

$$\alpha \leq (2\beta - 1)/(2(1 - \beta)).$$

This condition establishes an upper bound on parameter  $\alpha$ , which is decreasing with parameter  $\alpha$ .

offers as responders.

When inequity averse players are responders, they will only accept a certain proportion of the surplus and will reject any offer below this threshold level, depending on their degree of (disadvantageous) inequity aversion ( $\alpha$ ). We denote this proportion as their acceptance threshold ( $t^\alpha$ ). This share of the surplus is the result of making the responder player indifferent between accepting and rejecting the offer. In order to compute it, we equalize to zero the utility function where, without loss of generality, we have normalized the surplus to one. Thus  $t^\alpha - \alpha(1 - 2t^\alpha) = 0$  and therefore  $t^\alpha = \alpha/(1 + 2\alpha)$ . Note that this threshold is increasing in  $\alpha$  and strictly less than one-half for any finite  $\alpha$ .

If the inequity averse player gets to be the proposer in the ultimatum game it is easy to verify that it is a dominant strategy for him to always offer an equal split of the surplus. The fair offer will be accepted by his opponent whatever his type is. Notice that starting in an unequal distribution advantageous for him, giving an additional monetary unit to his opponent, reduces in one unit his material payoff and consequently his utility, but it reduces also in two units the inequity and as  $\beta > 0.5$ , increases his utility in more than one unit. The net effect is an increase in his total utility.

Summarizing, there is a dominant action for strongly inequity averse players, when they are proposers, which is to offer an equal split of the surplus and when they are responders, they will only accept a proportion of the current surplus greater or equal than  $t^\alpha$ . Notice that this result holds both under complete and incomplete information.

### 3.2 Behaviour of selfish players.

Selfish players accept any division of the surplus when they are responders in the ultimatum negotiation game, they will never reject any offer. However, the presence of a fraction of inequity averse agents in the population induces changes in the behaviour of selfish agents as proposers.

Let us analyze this behavior when a selfish player faces with probability 1 an inequity averse

player. The selfish player will anticipate that his opponent rejects any offer smaller than the threshold  $t^\alpha$ , and therefore, he will offer exactly a proportion  $t^\alpha$  of the surplus which will be accepted by the inequity averse player, although the latter gets an utility of zero.

If we assume that the row player is selfish and the column player is inequity averse, and solving by backward induction, the payoff matrix that the players face in the production stage is the following:

	C	NC
C	$(1 - t^\alpha + 1/2)B, B/2$	$d/2, d/2$
NC	$(1 - t^\alpha)d, 0$	$b, b$

It is easy to verify that cooperation is a dominant action for the selfish player. Thus, in this game with complete information, there is a unique perfect Nash equilibrium in which both types of players cooperate achieving the efficient outcome.

The intuition behind this result is the following. Cooperation is the best response for a selfish player also against non-cooperation, when he is confronted with probability one with an inequity averse opponent. The reason is that the latter does share equally even after defection. In other words, as strongly inequity averse players are very generous, selfish players do not fear being exploited by a defector.

However, players do not know the true type of the player with whom they are matched with in period  $t$ . They know, as was mentioned above, that there is a proportion  $p_t$  of self-interested agents and a proportion  $1 - p_t$  of inequity averse players.

The acceptance policy of a selfish player as a responder does not change with incomplete information. However, his behaviour as a proposer is indeed affected by the existence of a fraction of inequity averse individuals. Note that the particular realized surplus can also change his beliefs about his opponent's type. We will denote by  $\mu_t$  the updated probability that any player assigns to the fact that his opponent is selfish after observing the result of the production stage. Therefore, a selfish proposer has two options. First, to offer 0 and his expected payoff

would be the proportion  $\mu_t$  of the realized surplus because only selfish players would accept this offer. And, second, to offer the threshold level  $t^\alpha$  and his payoff would be a proportion  $(1 - t^\alpha)$  of the realized surplus, as in this case both types of players will accept. Therefore, offering 0 is better than offering  $t^\alpha$  when  $\mu_t \geq (1 - t^\alpha)$ .

The following lemma summarizes this result.

**Lemma 1** *If the selfish player gets to be the proposer in the negotiation stage, he will offer zero to his opponent if  $\mu_t \geq (1 - t^\alpha)$  and the acceptance threshold  $t^\alpha$  if  $\mu_t < (1 - t^\alpha)$ .*

Once we have analyzed the optimal bargaining behaviour of both types of players in the incomplete information game, we will next characterize the perfect bayesian equilibria of this game.

### 3.3 Equilibria with incomplete information.

The following two lemmas characterize the two pooling equilibria in which both types of players choose the same action in the production stage. We will set off-equilibrium beliefs to be as pessimistic as possible, that is, deviation are always believed to be by a selfish type.

**Lemma 2** *The Cooperative Equilibrium.*

*For every  $p_t$ , there exists a Perfect Bayesian Equilibrium in which both types of players choose the cooperative action  $C$  in the production stage of the game.*

*The equilibrium actions in the distribution stage are:*

1) *for inequity averse players: to offer half of the surplus ( $B$ ) as a proposer and to accept only offers greater or equal than  $t^\alpha 2B$  as a responder.*

2) *for selfish players: to accept every offer as a responder and to offer zero if  $p_t \geq (1 - t^\alpha)$  and to offer  $t^\alpha 2B$  if  $p_t < (1 - t^\alpha)$  as a proposer.*

Proof: see Appendix.



Notice that if  $p_t < (1 - t^\alpha)$  this cooperative equilibrium is efficient, and the equilibrium payoffs are  $p_t B + (1 - p_t)B/2$  and  $B - p_t B/2$  for the selfish player and the inequity averse player, respectively. However, if  $p_t \geq (1 - t^\alpha)$  this equilibrium is inefficient because inequity averse players reject the unfair offers of selfish players, leading to surplus destruction. In this case, the equilibrium payoffs are  $B - t^\alpha B + p_t t^\alpha B + B/2 - p_t B/2$  and  $B/2 + (1 - p_t)B/2$  for the selfish player and the inequity averse player respectively.

However, for some values of the distribution of preferences, there is also a pooling equilibrium in which no player chooses the cooperative action in the first stage, that is, the multiplicity of equilibria of the complete information stag-hunt game persists. In the next lemma we characterize this non-cooperative equilibrium.

**Lemma 3** *The Non-Cooperative Equilibrium.*

*If  $p_t \geq p' = (d - 2b)/d$ , there exists a Perfect Bayesian Equilibrium, in which both types of players choose non cooperation (NC) in the production stage.*

Proof: see Appendix.

The equilibrium payoff for both the selfish and the inequity averse player is  $b$ .

Next we comment on the existence of separating equilibria in which the two types of players choose different actions in the production stage. Firstly, non-cooperation by the selfish players and cooperation by the inequity averse players does not form part of an equilibrium. On the one hand, selfish players desire to separate from inequity averse players only if there is a sufficiently high proportion  $p_t$  of selfish players in the population. But, on the other hand, inequity averse players desire to separate from selfish players only if there is a sufficiently low proportion  $p_t$  of selfish players. It can be shown that there is no population preference distribution for which both incentive-compatible restrictions are satisfied simultaneously. Secondly, there is not either a separating equilibrium in which selfish players cooperate and inequity averse players do not, because the latter will deviate to cooperation. Concluding, there are no separating equilibria in this game. A formal proof of this statement can be found in the Appendix.

We can summarize the previous results in a more compact form in the following proposition:

**Proposition 1** *For every  $p_t$ , there exists a Perfect Bayesian Equilibrium in which both types of players choose the cooperative action in the production stage of the coordination game. Furthermore, if  $p_t \geq p' = (d - 2b)/d$ , there exists another Perfect Bayesian Equilibrium in which both types of players choose the non-cooperative action in the production stage.*

Given the results obtained in section 3.2 in a complete information scenario between a selfish and an inequity averse player, it is not surprising at all that for preferences distributions with a sufficiently low proportion of selfish players, i.e.  $p_t < p'$ , the unique Perfect Bayesian Equilibrium of the game with incomplete information is the Cooperative equilibrium. Cooperation is a dominant strategy for selfish players. Even if inequity averse players do not cooperate, as they are so generous in the negotiation stage offering half of the surplus  $d$ , selfish players prefer to choose cooperation instead of non-cooperation. Furthermore, if  $p_t < (1 - t^\alpha)$  selfish players will be also generous negotiating the surplus and the cooperative equilibrium would be efficient.

But for preferences distributions with a high proportion of selfish players, that is, for  $p_t > p'$ , there exists also a Non-cooperative equilibrium. The intuition is that, if there is a majority of selfish individuals in the population, then this type will not deviate from non-cooperation since they correctly expect that selfish players do not make fair offers in this situation. Concerning inequity averse players, if they deviate to cooperation, they will have to reject unfair offers of selfish players obtaining a payoff of zero.

The existence of multiplicity of equilibria for the range of parameters  $p_t \geq p'$ , poses a question on which will be the expected equilibrium a given period. This multiplicity problem will be tackled in a dynamic framework that we will introduce in the next section.

## 4 The socialization process and the parent's educational effort.

In this section we present the dynamic model that describes the evolution of the distribution of preferences along with strategies in the two-stage coordination game analyzed in the previous sections.

We consider overlapping generations of agents who only live two periods (as a young and as an adult). In the first period, the agent is a child and is educated in certain preferences and in the second period (as an adult with well defined preferences), is randomly matched with another adult player, to play the two-stage coordination game. In this second period, any adult player has one offspring and has to make a (costly) decision regarding his child's education, trying to transmit his own preferences. As it is usual in this sort of model we assume that fertility is exogenous, that is, an adult has only one child independently of his performance in the two-stage game, and thus the population size remains constant. It is also assumed that reproduction is asexual, with a parent per child.

Preferences among players are influenced by a purposeful and costly socialization process. Children acquire preferences through observation, imitation and learning of cultural models prevailing in their social and cultural environment. We will draw from the model of cultural transmission of Cavalli-Sforza and Feldman (1981) and, in particular, from Bisin and Verdier (1998, 2000).

Let  $\tau^i \in [0, 1]$  be the educational effort made by a parent of type  $i$  where  $i \in \{e, a\}$  and  $e$  denotes selfish and  $a$  denotes strongly inequity averse.

The socialization mechanism works as follows. Consider a parent with  $i$  preferences. His child is first directly exposed to the parent's preferences and is socialized to this preferences with probability  $\tau^i$  chosen by the parent (vertical transmission); with complementary probability  $1 - \tau^i$ , this direct socialization fails, and the child is socialized to his preferences by a randomly chosen member of the population (oblique transmission). In this case, his preferences are selfish with probability  $p_t$  and inequity averse with probability  $(1 - p_t)$ .

Let  $P^{ij}$  denote the probability that a child of a parent with preferences  $i$  is socialized to preferences  $j$ . The socialization mechanism is then characterized by the following transition probabilities:

$$P_t^{ee} = \tau_t^e + (1 - \tau_t^e)p_t \quad (2)$$

$$P_t^{ea} = (1 - \tau_t^e)(1 - p_t) \quad (3)$$

$$P_t^{aa} = \tau_t^a + (1 - \tau_t^a)(1 - p_t) \quad (4)$$

$$P_t^{ae} = (1 - \tau_t^a)p_t \quad (5)$$

Given these transition probabilities it is easy to characterize the dynamic behavior of  $p_t$ :

$$p_{t+1} = [p_t P_t^{ee} + (1 - p_t) P_t^{ae}]. \quad (6)$$

Substituting (2) to (5) we obtain:

$$p_{t+1} = p_t + p_t(1 - p_t)[\tau_t^e - \tau_t^a]. \quad (7)$$

Direct transmission is justified because parents are altruistic towards their children. But, an important feature is that they have some kind of imperfect altruism: their socialization decisions are not based on the purely material payoff expected for their children but on the payoff as perceived by their parents according to their own preferences. This particular form of myopia is called imperfect empathy. As a consequence, the cultural dynamics is not necessarily payoff-monotonic. Direct transmission is also costly. Let  $C(\tau^i)$  denote the cost of the education effort  $\tau^i$ ,  $i \in \{e, a\}$ . While it is possible to obtain similar results with any increasing and convex cost function we will assume, for simplicity, the following quadratic form  $C(\tau^i) = (\tau^i)^2/2k$ , with  $k > 0$ . Therefore, a parent of type  $i$  chooses the education effort  $\tau^i \in [0, 1]$  at time  $t$ , which maximizes

$$P_t^{ii}(\tau^i, p_t) V^{ii}(p_{t+1}^E) + P_t^{ij}(\tau^i, p_t) V^{ij}(p_{t+1}^E) - (\tau^i)^2/2k,$$

where  $P^{ij}$  are the transition probabilities and  $V^{ij}$  is the utility to a parent with preferences  $i$  if his child is of type  $j$ . Notice that the utility  $V^{ij}$  depends on  $p_{t+1}^E$ , which denotes the expectation about the proportion of selfish players in period  $t + 1$ . In this work we will assume that parents

have adaptive or backward looking expectations, believing that the proportion of selfish players will be the same in the next period as in the current period, that is,  $p_{t+1}^E = p_t$ .

Maximizing the above expression with respect to  $\tau^i$ ,  $i \in \{e, a\}$ , we get the following optimal education effort functions<sup>4</sup>:

$$\tau^{e*}(p_t) = k \cdot \Delta V^e(p_t) \cdot (1 - p_t),$$

$$\tau^{a*}(p_t) = k \cdot \Delta V^a(p_t) \cdot p_t.$$

Here  $\Delta V^e = V^{ee} - V^{ea}$  and  $\Delta V^a = V^{aa} - V^{ae}$ . That is,  $\Delta V^i$  is the net gain from socializing your child to your own preferences. It can also be interpreted as the cultural intolerance of parents with respect to cultural deviation from their own preferences. According to the imperfect empathy notion, parents obtain a higher utility if their children share their preferences, so these levels of cultural intolerance are non-negative.

Note that the education effort  $\tau^e(p_t)$  of a selfish parent decreases with the proportion of selfish individuals in the population. The reason is very intuitive: the larger  $p_t$  is, the better children are socialized to the selfish preferences in the social environment. On the contrary, the educational effort chosen by the inequity averse players  $\tau^a(p_t)$  increases with  $p_t$ , that is, the greater the proportion of selfish players in the population, the bigger the socialization effort of the former parents in order to offset the pressure of the environment if they want their children to share their own preferences. In other words, oblique transmission acts as a substitute for vertical transmission. Bisin and Verdier (2000) refer to this feature of educational effort as the cultural substitution property.

The other determinant of the optimal education effort is the relative profit  $\Delta V^i$  to a parent of type  $i$  from transmitting her own cultural traits. Its value will depend on the expected equilibrium of the game that their children, as adults, will face in the next period.

---

<sup>4</sup>In order to have interior solutions the parameter  $k$  must be chosen small enough so that in equilibrium  $\tau^i < 1$ .

## 5 Migration, preference distribution and strategies in the long run.

In this section we will characterize the steady states of the cultural dynamics when the individuals coordinate in an equilibrium of the stag-hunt game. However, in section 3 we have shown that for initial preference distributions greater than  $p'$ , there are two possible equilibria in the society depending on whether the population coordinates in the cooperative equilibrium or in the non-cooperative equilibrium. Therefore, for these values of the parameters we still have an equilibria multiplicity problem.

Some authors have found this same problem and have just decided to select one of these equilibria, that is, they assume that all the population or groups coordinate in the same equilibrium. See, for instance, Bisin, Topa and Verdier (2004) and Mengel (2008). We think that this approach leaves unsolved the equilibrium selection problem. Instead of that, in order to obtain endogenous equilibrium selection we will analyze the effects on the dynamics of the possibility of migration among different groups.

We will assume that there are several populations (groups) where each one expects and plays a particular equilibrium of the coordination game. Socialization takes place inside the group but there is a (probably small) rate of migration  $\gamma$  among groups. That is, a proportion  $\gamma$  of adult individuals migrate to another group to play the coordination game. They will play according to the expectations and customs of their new group. In this new scenario, parents in each group will socialize their offspring taking now into account that when their children become adults, they will have a positive probability of migrating and therefore ending up playing a different equilibrium.

In order to simplify the analysis, we will start analyzing this situation in absence of migration ( $\gamma = 0$ ) and we will characterize the steady states of a society in which there are isolated populations and each population of individuals coordinates in an equilibrium of the stag-hunt game. We need to know the optimal level of educational effort that parents in each population

chooses in each period. This depends on the net gains for parents from transmitting their own preferences that are given by expression  $V^{ij}$ .

### 5.1 The long-run distribution of preferences in a cooperative population.

We analyze the distribution of preferences in a population where parents play and expect to be played the cooperative equilibrium. Notice that it might happen for any  $p_t$ . We denote this population as a cooperative population.

Firstly, we calculate the net gains for parents from transmitting their own preferences  $\Delta V^i$ , that is, their levels of cultural intolerance, when  $p_t \geq 1 - t^\alpha$ . In this case, selfish players will offer zero as proposers in the negotiation and the inequity averse players will reject the offer.

Recall that to compute  $V^{ij}$  we suppose that a parent of type  $i$  evaluates his child's well-being using his own utility function. For example, when  $p_t \geq 1 - t^\alpha$ ,  $V^{ae}$  is the utility to an inequity averse parent if his child is selfish. This child, as adult, will offer nothing to his opponent when he is a proposer and will accept any offer as a responder. This behavior produces inequality in the distribution of the payoffs for his child (advantageous in the first case and disadvantageous in the second). This inequality reduces the utility in his parent's eyes. Note that although the expected material payoff for his child is  $p_t B + (1 - p_t)B/2$ , the utility for the inequity averse parent is  $B/2(1 + p_t) - p_t\beta B - p_t\alpha B$ .

It can be verified that for selfish parents:

$$V^{ee} = p_t B + (1 - p_t)B/2, \text{ and}$$

$$V^{ea} = B/2 + (1 - p_t)B/2.$$

Therefore, the level of cultural intolerance of a selfish parent is:  $\Delta V^e = V^{ee} - V^{ea} = (p_t - \frac{1}{2})B \geq 0$ .

It can be checked also that for inequity averse parents:

$$V^{aa} = B/2 + (1 - p_t)B/2, \text{ and}$$

$$V^{ae} = B/2(1 + p_t) - p_t\beta B - p_t\alpha B.$$

Thus, the level of cultural intolerance of an inequity averse parent is:  $\Delta V^a = V^{aa} - V^{ae} = B(1/2 + p_t(\alpha + \beta - 1)) \geq 0$ .

Notice that both types of parents have strictly positive levels of cultural intolerance and consequently, they will have incentives for active socialization.

From the previous expressions and from section 3 we can obtain the optimal educational effort functions for both types of players when  $p_t \geq 1 - t^\alpha$ , which are:

$$\begin{aligned}\tau_t^{*e} &= k(p_t - \frac{1}{2})B(1 - p_t), \\ \tau_t^{*a} &= kB(1/2 + p_t(\alpha + \beta - 1))p_t.\end{aligned}$$

Secondly, we compute the levels of cultural intolerance  $\Delta V^i$  when  $p_t < 1 - t^\alpha$ . Under this condition selfish players offer the threshold which is accepted by both types of players. Recall that the inequity averse type obtains a positive material payoff by accepting but an utility of zero.

It can also be verified that for selfish parents:

$$\begin{aligned}V^{ee} &= (1 - t^\alpha)B + p_t t^\alpha B + (1 - p_t)B/2 \text{ and} \\ V^{ea} &= B/2 + p_t t^\alpha B + (1 - p_t)B/2.\end{aligned}$$

Thus, the level of cultural intolerance of a selfish parent is:

$$\Delta V^e = V^{ee} - V^{ea} = B(1/2 - t^\alpha) \geq 0.$$

For inequity averse parents:

$$\begin{aligned}V^{aa} &= B/2 + (1 - p_t)B/2 \text{ and} \\ V^{ae} &= B - t^\alpha B - \beta(B(1 - 2t^\alpha)) + p_t[(t^\alpha B - \alpha(B(1 - 2t^\alpha)))] + (1 - p_t)B/2.\end{aligned}$$

Therefore, the level of cultural intolerance of an inequity averse parent is:

$$\Delta V^a = V^{aa} - V^{ae} = B(t^\alpha + \beta(1 - 2t^\alpha) - 1/2) \geq 0.$$

Then, if  $p_t < 1 - t^\alpha$ ,

$$\begin{aligned}\tau_t^{*e} &= kB(1/2 - t^\alpha)(1 - p_t), \\ \tau_t^{*a} &= kB(t^\alpha + \beta(1 - 2t^\alpha) - 1/2)p_t.\end{aligned}$$

Substituting the previous optimal education effort functions in the equation in differences of section 4 we get the following two-branch dynamics under the assumption of backward looking



expectations:

**Dynamics A:**

$$p_{t+1} = p_t + p_t(1 - p_t)[-kBp_t^2 + kB(3/2)p_t - kB/2 - (kp_tB/2 + kBp_t^2(\alpha + \beta - 1))]$$

if  $p_t \geq 1 - t^\alpha$ .

**Dynamics B:**

$$p_{t+1} = p_t + p_t(1 - p_t)[Bk(1/2 - t^\alpha - p_t/2 + p_t t^\alpha) - (kB(t^\alpha + \beta(1 - 2t^\alpha) - 1/2)p_t)]$$

if  $p_t < 1 - t^\alpha$ .

Notice that there is a discontinuity in  $p_t = 1 - t^\alpha = (1 + \alpha)/(1 + 2\alpha)$ . The phase diagram in Figure 1 shows this case.

(Insert Figure1 here)

The following proposition characterizes the globally stable steady-state of this population.

**Proposition 2** *For any  $p_t \in (0, 1)$ , the preference distribution of a cooperative population converges to  $p^* = 1/(2\beta)$ , where  $p^*$  is such that  $\tau_t^a(p^*) = \tau_t^e(p^*)$  in dynamics B. In this steady state the efficient cooperative equilibrium is played.*

Proof: see Appendix.

The complete and formal analysis of this result is relegated to the appendix, but let us provide some intuition on this proposition.

Note that the dynamics has the following steady states:  $p = 0$ ,  $p = 1$  and the interior steady state  $p^* = 1/(2\beta)$ , where  $p^*$  is such that both educational effort levels get equalized under dynamics (B).

The steady state  $p = 0$  is unstable. This steady state is a completely homogeneous distribution of preferences, that is, all players being inequity averse. If the selfish players are in a minority (that is,  $p$  is very close to 0), their socialization effort will be very intensive in an attempt to offset the effect of oblique transmission. In this context,  $\tau^{*e}$  exceeds  $\tau^{*a}$  and the selfish preferences will spread over generations preserving their presence in the society. A similar argument explains why  $p = 1$  is also unstable.

On the other hand, in the dynamics A,  $\tau_t^{*a}$  is always greater than  $\tau_t^{*e}$ ,  $\forall p_t \in (0, 1)$ . This implies that the trajectory of dynamics A is always decreasing in the range  $(1 - t^\alpha, 1)$ .

Therefore, the inefficient cooperative equilibrium, in which there is surplus destruction, can not be a steady state of the cultural dynamics. The intuition of this result is that when the proportion of selfish players is high, i.e. for  $p_t > (1 - t^\alpha)$ , selfish players follow a very greedy strategy when they have all the bargaining power (i.e. as proposers). This yields a very high level of cultural intolerance of inequity averse parents with respect to cultural deviation from their own preferences  $\Delta V^a$ , as compared with the level of cultural intolerance of the selfish parents  $\Delta V^e$ . Consequently, the socialization effort of inequity averse parents is greater than that of selfish parents, leading to a reduction of the proportion of selfish players  $p_t$  over generations.

Eventually, the dynamics will reach the region where  $p_t < (1 - t^\alpha)$ . Then selfish players will offer the threshold when they are proposers because of the presence of a significant fraction of inequity averse players who will punish with rejection any greedy offer. In this branch B of the dynamics there is an interior rest point  $p^*$  which is the globally stable steady state. In this steady state it is played the efficient cooperative equilibrium in which there is no surplus destruction.

## 5.2 The long-run distribution of preferences in a non-cooperative population.

Suppose now an isolated population where parents play and expect to be played the non-cooperative equilibrium. This can only happen if  $p_t \geq p^* = (d - 2b)/d$ . Denote this population as a non-cooperative population.

Notice that  $V^{ij} = b$  for all type of parents. Therefore, the net gains for any type of parent from transmitting their own preferences  $\Delta V^i$ , that is, their levels of cultural intolerance are zero. But this, in turn, implies that there are not incentives at all for socialization for any parent. Therefore, all the optimal education effort functions are zero for both types and consequently,

in a non-cooperative group there are not incentives for socialization and the distribution of preferences will remain unchanged, that is,  $p_{t+1} = p_t$ .

Let us summarize this result in the next proposition:

**Proposition 3** *If  $p_t \geq p' = (d - 2b)/d$ , then in a non-cooperative population any initial distribution  $p_t \geq p'$  is a stable stationary state.*

( Insert Figure 2 here)

Thus, if a population coordinates in the non-cooperative equilibrium, it will remain locked in the same distribution of preferences. The phase diagram in Figure 2 shows this case.

### 5.3 Migration between populations and cooperation.

In this section we will allow adult members of different populations (groups) to migrate to other population to play the coordination game. We assume that socialization takes place inside the population but there is a (probably small) rate of migration  $\gamma$  among populations. They will play according to the expectations and customs of their new population. But in this new situation, parents will socialize their children taking into account that when the latter become adults they will have a positive probability of migrating and therefore playing a different equilibrium of the group in which the children were raised.

Denote as  $p_o^j$  the initial distribution of preferences in a population  $j$ . For sake of simplicity, we will suppose that there are only two groups or populations in the society.

If  $p_o^j < p'$ , then this population will coordinate in the unique perfect bayesian equilibrium for this preferences distribution: the cooperative equilibrium.

But for populations such that  $p_o^j > p'$ , since there exists multiplicity of equilibria, more possibilities arise. In order to discard the extreme cases in which all groups play the same equilibrium, we will make the following *Population Heterogeneity Assumption*: if both populations have  $p_o^j > p'$ , then we will assume that in at least one population it is played the cooperative

equilibrium. While if in one population holds  $p_o^j < p'$  and in the other  $p_o^j > p'$ , we will assume that in the latter it is played the non-cooperative equilibrium.

Let us also assume that  $p' > p^*$  (which is the case for payoff  $d$  being very large in comparison to payoff  $b$ ). The next proposition states the main result of our paper.

**Proposition 4** *Suppose, without loss of generality, that there are two populations with initial preference distributions,  $p^1$  and  $p^2$ , and assume that both the assumption of population heterogeneity and  $p' > p^*$  hold, then for any  $p^1, p^2 \in (0, 1)$  both populations converge to a preference distribution  $p^* = 1/(2\beta)$ , where the efficient cooperative equilibrium is played.*

Proof: see Appendix<sup>5</sup>.

The intuition of this result is the following. Let us assume that one population is playing the cooperative equilibrium and the other population is initially playing the non-cooperative equilibrium. Denote as  $p^{nc}$  the proportion of selfish players in the latter group. Notice that, at least initially  $p^{nc} > p'$ .

Suppose that we are in the cooperative population. In this case, the possibility of migration to the non-cooperative population slows the trajectory to the steady state, but does not change it qualitatively.

Suppose now that we are in the population where the non-cooperative equilibrium is played. Parents in this group have now incentives to socialize their offspring due to the bigger utility that their children can obtain by migrating to the cooperative group.

Moreover, the educational effort of inequity averse parents is greater than the effort of selfish parents, given the relatively high proportion of selfish individuals ( $p^{nc}$ ) in this non cooperative population. The educational effort might also depend indirectly on the proportion of selfish individuals in the cooperative population ( $p^c$ ) where their children migrate into. In particular, it depends on both the levels of cultural intolerance  $\Delta V^i$  and the dynamics followed in the cooperative population. If the latter follows dynamics B, these levels are independent of the population

---

<sup>5</sup>We are very grateful to a referee for providing a better and more concise alternative proof for this proposition.

distribution of selfish players in the cooperative group. However, if it follows dynamics A, the degree of cultural intolerance depends on  $p^c$ , so the dynamics of the preferences distribution of the non-cooperative group depends both on the proportions of selfish individuals in the home population ( $p^{nc}$ ) and in the cooperative population ( $p^c$ ). But recall that the incentives for socialization under dynamics A are always greater for the inequity averse parents than for the selfish parents as we have explained in the previous section. Therefore, as a consequence of both the direct and the possible indirect effect, the proportion of selfish players of the non-cooperative group ( $p^{nc}$ ) decreases over time and eventually reaches a value smaller than  $p'$ . At this point non-cooperation is not an equilibrium and this population switches to the unique cooperative equilibrium.

Note that if  $p' < p^*$ , the result stated in the proposition is not always valid. It does not hold for any initial condition of the dynamics, since it might happen that the proportion of selfish individuals in the non-cooperative group ( $p^{nc}$ ) never gets smaller than the critical value  $p'$ . More precisely, it can be proved that whatever the relationship between  $p'$  and  $p^*$ , both populations are going to converge to  $p^*$ . However, if  $p' \leq p^*$ , then, by assumption, one of the populations must be playing non-cooperatively. So, migration in this case is not a sufficient condition for the spread of cooperation between groups.

## 6 Conclusions.

Cultural transmission plays an important role in the formation of many preference traits and norms, like attitudes towards family, in the job market and cooperation. The aim of this paper is to analyze if cooperation can be the product of cultural evolution in a stag hunt game. Most of the economic interactions that turn out to have a coordination game structure are actually sequential two-stage games. Players firstly make productive decisions, generating a joint surplus, and subsequently, they bargain over its division among them. Moreover, if the players' choices made in the production stage determine the distribution of bargaining power in the second

distributive stage, this stag hunt game is known as the assurance or mutualism game.

We have presented an overlapping generations model with cultural transmission of preferences where the distribution of preferences in the population and the strategies in the two-stage coordination game are determined endogenously and simultaneously. We assume that there are several groups in the society, some of them play cooperatively and others do not. Socialization takes place inside the group, but there is a positive rate of migration among groups which parents anticipate. We show how all groups converge to the cooperative equilibrium.

If the shares in the division of the surplus are predetermined, that is, not the consequence of negotiation with different previous-choice-dependent bargaining powers, then the two-stage game results in a coordination game only if  $d \leq 2b$ . In this class of stag hunt games migration and cultural transmission do not yield global convergence to the cooperative equilibrium. The reason is that non-cooperation is an equilibrium for any preference distribution, even with a very high proportion of other-regarding individuals. Which additional features of cultural transmission, if any, can solve convergence to cooperation in this latter class of stag hunt games belongs to future research.

A natural future extension of our work is to explore the effects of different types of influences among groups (other than migration), as for instance, the existence of a degree of interaction in oblique transmission. We will also consider the effects of other biased cultural mechanisms such as conformism or prestige-based imitation.

## APPENDIX.

### Proof of lemma 2.

We will show that there is not unilateral profitable deviation for any type of player if players pool in cooperation (C) in the first stage. Suppose first that  $p_t \geq (1 - t^\alpha)$ . Selfish players, as proposers, will offer 0 to their opponents and will accept any offer when they are responders.

The expected payoff of choosing C for the selfish type will be  $p_t B + (1 - p_t)B/2$ .

If this type of player deviates to NC, he will obtain an expected payoff of  $p_t d$ . As it can be checked this type of player will not deviate.

On the other hand, a strongly inequity averse type of player will offer half of the surplus when she is the proposer, and this offer will be accepted by both types of players, obtaining an expected payoff of  $(1/2)B$ . But when she is the responder, she will only accept the threshold level and as the selfish player will offer nothing. So she will reject the offer and both players will have a payoff of 0 with probability  $p_t$ , but with probability  $(1 - p_t)$  the inequity averse opponent will offer  $B$  to her. In summary, the expected payoff of an inequity averse player of choosing C will be  $(1/2)B + (1/2)(1 - p_t)B = p_t B/2 + (1 - p_t)B$ .

Her expected payoff of deviating to NC will be  $d/2$ . Therefore, choosing C is also a best response for this type of player.

Suppose now that  $p_t < (1 - t^\alpha)$ . In this case selfish players offer  $t^\alpha 2B$  when they are proposers. The expected payoff of choosing C for a selfish type will be  $p_t B + (1 - p_t)(1 - t^\alpha + 1/2)B$ .

If he deviates to NC his expected payoff will be  $(1 - t^\alpha)d$ . As  $(1 - t^\alpha) \in (1/2, 3/4)$ , then this player will not deviate.

The utility of inequity averse players of playing C in this case will be  $B/2 + (1 - p_t)B/2$ .

A deviation to NC yields a payoff of  $(d/2)$  smaller than that from C.

### Proof of lemma 3.

Let us check under which conditions there is an equilibrium where players choose the non cooperative action in the production stage of the stag hunt game. In this pool both types of players obtain a payoff of  $b$ .

Suppose that players believe that a deviation to cooperation comes from a selfish player. Then, if a selfish player deviates by choosing C, he will obtain a payoff of 0 when matched with a selfish opponent and  $(d/2)$  when matched with an inequity averse opponent. Therefore, the expected payoff of the deviation is  $(1 - p_t)d/2$ . Thus, the selfish players will not deviate from choosing NC if  $b \geq (1 - p_t)d/2$ , that is, if  $p_t \geq (d - 2b)/d$ .

Notice that exactly the same argument applies for the inequity averse type under the proposed off-equilibrium beliefs.

**Proof that there is not a separating equilibrium in the two-stage game with incomplete information.**

Let us firstly check that selfish players choosing C and inequity averse players choosing NC is not an equilibrium.

The expected payoff for an inequity averse player in this strategy combination is  $p_t(d/2) + (1 - p_t)b$ . Deviating to C and given that players expect C from a selfish player, the inequity averse deviator obtains a payoff of  $p_t B/2 + (1 - p_t)(d/2)$ . So the inequity averse type deviates imitating the selfish type.

Consider now a situation in which selfish players choose NC and inequity averse players choose C. Let us check if this strategy profile constitutes an equilibrium.

The expected payoff for a selfish player in this profile is  $p_t b + (1 - p_t)(1 - t^\alpha)d$  and for an inequity averse player is  $(1 - p_t)B$ .

The selfish type imitating the inequity averse player gets an expected payoff of  $p_t(t^\alpha d) + (1 - p_t)(1 - t^\alpha + \frac{1}{2})B$ . Therefore, the incentive constraint for the selfish player holds if  $p_t \geq p^a(t^\alpha) = \frac{B(1-t^\alpha+1/2)-(1-t^\alpha)d}{B(1-t^\alpha+1/2)-(d-b)}$ .

On the other hand, the inequity averse type imitating the selfish type gets an expected payoff of  $p_t b + (1 - p_t)(d/2)$ . Therefore, the incentive constraint for the inequity averse type holds if  $p_t \leq p^b = \frac{B-d/2}{B-d/2+b}$ .

But  $p^a(t^\alpha) > p^b$  for all  $t^\alpha \in (1/4, 1/2)$ . Notice that this inequality holds for  $t^\alpha = 1/4$  and for  $t^\alpha = 1/2$ , and the function  $f(t^\alpha) = p^a(t^\alpha) - p^b$  is monotone.



Therefore, we conclude that the set of preference distributions for which the incentive constraints of both types hold simultaneously is empty.

**Proof of proposition 2.**

According to dynamics A,  $\tau_t^a$  is always greater than  $\tau_t^e$ ,  $\forall p_t \in (0, 1)$ . This is because the equation:  $(\alpha + \beta)p_t^2 - p_t + 1/2 = 0$ , has no real roots.

This implies that the behavior of dynamics (A) is always decreasing in the range  $(1 - t^\alpha, 1)$  and eventually  $p_t$  will fall below  $(1 - t^\alpha)$ .

Under assumption (2),  $p^* = 1/(2\beta) \leq 1 - t^\alpha$ , then dynamics (B), defined in the interval  $[0, 1 - t^\alpha)$ , has an homogeneous steady state  $p = 0$  which is unstable and an interior steady state  $p = p^* = 1/(2\beta)$  which is globally stable, where  $p^*$  is such that  $\tau^{*e} = \tau^{*a}$ . (See, Bisin and Verdier, 2001, Proposition 1).

**Proof of proposition 4.**

First, we will introduce some additional notation. Denote  $\Delta V_j^i$  the level of cultural intolerance of a parent of type  $i$  when he is in population  $j$  and his child has a probability  $\gamma$  of migrating to the other population  $-j$ , where  $i = e, a$  and  $j = 1, 2$ . Therefore,  $\Delta V_j^i = (1 - \gamma)\Delta V_j^i + \gamma\Delta V_{-j}^i$ .

Let  $\tau_j^i$  denote the educational effort that type  $i$  in population  $j$  would exert if population  $j$  were in isolation.

We define  $\tau_j^i$  as the effort put into cultural transmission by an agent of type  $i$  in population  $j$  when that population is not in isolation. Note that  $\tau_j^e = k \Delta V_j^e(1 - p^j)$  and  $\tau_j^a = k \Delta V_j^a p^j$ . Then,

$$\begin{aligned}\tau_j^e &= (1 - \gamma) \tau_j^e + \gamma \tau_{-j}^e \left( \frac{1-p^j}{1-p^{-j}} \right) \\ \tau_j^a &= (1 - \gamma) \tau_j^a + \gamma \tau_{-j}^a \left( \frac{p^j}{p^{-j}} \right)\end{aligned}$$

We know from the dynamic behavior of  $p^j$  (equation (7)) that  $p_{t+1}^j - p_t^j$  and  $\tau_j^e - \tau_j^a$  have the same sign. Then we can compute:

$$\tau_j^e - \tau_j^a = (1 - \gamma) (\tau_j^e - \tau_j^a) + \gamma [\tau_{-j}^e \left( \frac{1-p^j}{1-p^{-j}} \right) - \tau_{-j}^a \left( \frac{p^j}{p^{-j}} \right)] \quad (A)$$

We are interested in the direction and the convergence of the dynamics, that is, whether  $p^j$  is decreasing or decreasing and whether it converges to  $p^*$ .

Let us assume that population 1 is playing the cooperative equilibrium and population 2 is initially playing the non-cooperative equilibrium, that is,  $p^2 \geq p'$ . Then we have

$$\tau_1^e - \tau_1^a = (1 - \gamma) (\tau_1^e - \tau_1^a) \quad (\text{B})$$

$$\tau_2^e - \tau_2^a = \gamma [\tau_1^e(\frac{1-p^2}{1-p^1}) - \tau_1^a(\frac{p^2}{p^1})] \quad (\text{C}).$$

Clearly as long as  $p^2 \geq p'$ , population 1 is going to behave as if it were in isolation.

Hence, as long as  $p^2 \geq p'$ , we have that  $p^1$  will tend to  $p^*$ , independently if  $p^1$  is greater or smaller than  $p'$ . One of two things can happen. If  $p^1$  converges to  $p^*$  while  $p^2 > p^*$ , then  $p^1 = p^* < p' \leq p^2$ . If we look at equation C, we see that  $p^2$  is decreasing. Consequently, it must eventually be the case that  $p^2 < p'$ . At this point, both populations are playing the cooperative equilibrium. The other possibility is that before  $p^1$  converges to  $p^*$ ,  $p^2$  becomes smaller than  $p'$ . At this point both populations are playing also the cooperative equilibrium.

Therefore, it only remains to be proved convergence to  $p^*$  when both populations play the cooperative equilibrium in the presence of migration. Assume without loss of generality that  $p^1 < p^2$ .

If both  $p^1, p^2 < p^* < p'$  (or  $p' > p^1, p^2 > p^*$ ) then it is easy to check with equation A that both populations converge to  $p^*$ .

Problems might arise when  $p^1 < p^* < p^2 < p'$ . We will denote by  $\Delta V_j^i(B)$ , the level of cultural intolerance of parents in a cooperative group if it follows dynamics B and  $\Delta V_j^i(A)$ , if it follows dynamics A. Recall that in the first case this level is a constant and in the second one depends on the proportion  $p^j$ .

Assume first that both populations follow dynamics B, that is,  $p^2 < 1 - t^\alpha$ . Then substituting in equation A,

$$\tau_j^e - \tau_j^a = k (\Delta V_j^e(B)(1 - p^j) - \Delta V_j^a(B)p^j).$$

This expression is positive for  $p^1 < p^*$  and it is negative for  $p^2 > p^*$ . Therefore both,  $p^1$  and  $p^2$  tend to  $p^*$

Now consider that  $p' > p^2 > 1 - t^\alpha$ . Hence, population 1 follows dynamics B and population 2 follows initially dynamics A. The sign of  $\tau_1^e - \tau_1^a$  is at first glance undetermined because the

first term of equation A is positive but the second is negative. However,  $\underline{\tau}_2^e - \underline{\tau}_2^a$  is negative, substituting in equation A,

$$\underline{\tau}_2^e - \underline{\tau}_2^a = (1 - \gamma)k(\Delta V_2^e(A)(1 - p^2) - \Delta V_2^a(A)p^2) + \gamma k(\Delta V_2^e(B)(1 - p^2) - \Delta V_2^a(B)p^2).$$

We know that in the dynamics A,  $\tau^a$  is always greater than  $\tau^e$ ,  $\forall p_t \in (0, 1)$ . Therefore, the first term is negative, but the second term is also negative because  $p^2 > p^*$ . Consequently,  $p^2$  is decreasing and it must eventually be the case that  $p^2 < 1 - t^a$ . At this point both populations are following dynamics B.

For completeness, let us show convergence to  $p^*$  for both populations when  $p' < p^*$ . Of course, if both  $p^j$  are less than  $p' < p^*$ , then they will both tend to increase. Hence, we must eventually arrive at the case where at least one  $p^j > p'$ . So in this case, let us suppose that population 1 plays cooperatively, and population 2 plays non-cooperatively. As before,  $p^1$  will tend to  $p^*$ . If  $p^* \approx p^1 < p^2$ , then  $p^2$ , by dynamics B or A, will decrease to  $p^*$  and if  $p^* \approx p^1 > p^2$ , then  $p^2$ , by dynamics B, will increase to  $p^*$ .

So we can see that whatever the relationship between  $p'$  and  $p^*$ , both populations are going to converge to  $p^*$ . However, if  $p' \leq p^*$ , then, by assumption, one of the populations must be playing non-cooperatively.

## REFERENCES.

- Bisin, A., Verdier, T., 1998. On the Cultural Transmission of Preferences for Social Status. *Journal of Public Economics* 70, 75-97.
- Bisin, A., Verdier, T., 2000. Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits. *Quarterly Journal of Economics* 115, 955-988.
- Bisin, A., Verdier, T., 2001. The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory* 97, 298-319.
- Bisin, A., Topa, G., Verdier, T., 2004. Cooperation as a Transmitted Cultural Trait. *Rationality and Society* 16, 477-507.

- Bolton G., Ockenfels A., 2000. ERC: A Theory of Equity, Reciprocity and Competition. *American Economic Review* 90, 166-193.
- Boyd, R., Richerson, P., 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- Cavalli-Sforza, L., Feldman, M., 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton NJ: Princeton University Press.
- Charness, G., Rabin, M., 2002. Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics* 117, 817-869.
- Cooper, R., 1999. *Coordination Games: Complementarities and Macroeconomics*. New York: Cambridge University Press.
- Dawid, H., MacLeod, W.B., 2001. Hold Up and the Evolution of Bargaining Conventions. *European Journal of Economics and Social Systems* 15, 153-69.
- Dufwenberg, M., Kirchsteiger, G., 2004. A Theory of Sequential Reciprocity. *Games and Economic Behavior* 47, 268-98.
- Ellingsen, T., Robles, J., 2002. Does Evolution Solve the Hold Up Problem? *Games and Economic Behavior* 39, 28-53.
- Ellison, G., 1993. Learning, Local Interaction, and Coordination. *Econometrica* 61, 1047-1071.
- Fehr, E., Schmidt, K., 1999. A Theory of Fairness, Competition and Cooperation. *Quarterly Journal of Economics* 114, 817-868.
- Kandori, M., Mailath, G., Rob, R., 1993. Learning, Mutation and Long Run Equilibria in Games. *Econometrica* 61, 29-56.
- Kuran, T., Sandholm, W., 2008. Cultural Integration and Its Discontents. *Review of Economic Studies* 75, 201-228.
- Mengel, F., 2008. Matching Structure and the Cultural Transmission of Social Norms. *Journal of Economic Behavior and Organization* 67, 608-623.
- Olcina, G., Penyarrubia, C., 2002. Specific Investments and Coordination Failures. *Eco-*

nomics Bulletin 3, 1-7.

Rabin, M., 1993. Incorporating Fairness into Game Theory and Economics. *American Economic Review* 83, 1281-1302.

Sandholm, W., 2001. Preference Evolution, Two-Speed Dynamics, and Rapid Social Change. *Review of Economic Dynamics* 4 , 637-679.

Trogër, T., 2002. Why Sunk Costs Matter for Bargaining Outcomes: An Evolutionary Approach. *Journal of Economic Theory* 102, 375-402.

Van Huyck, J., Battalio, R., Beil, R., 1990. Tacit Coordination Games, Strategic Uncertainty and Coordination Failure. *American Economic Review* 80, 234-248.

Young, H. P., 1993. The Evolution of Conventions. *Econometrica* 61, 57-84.

Figure1

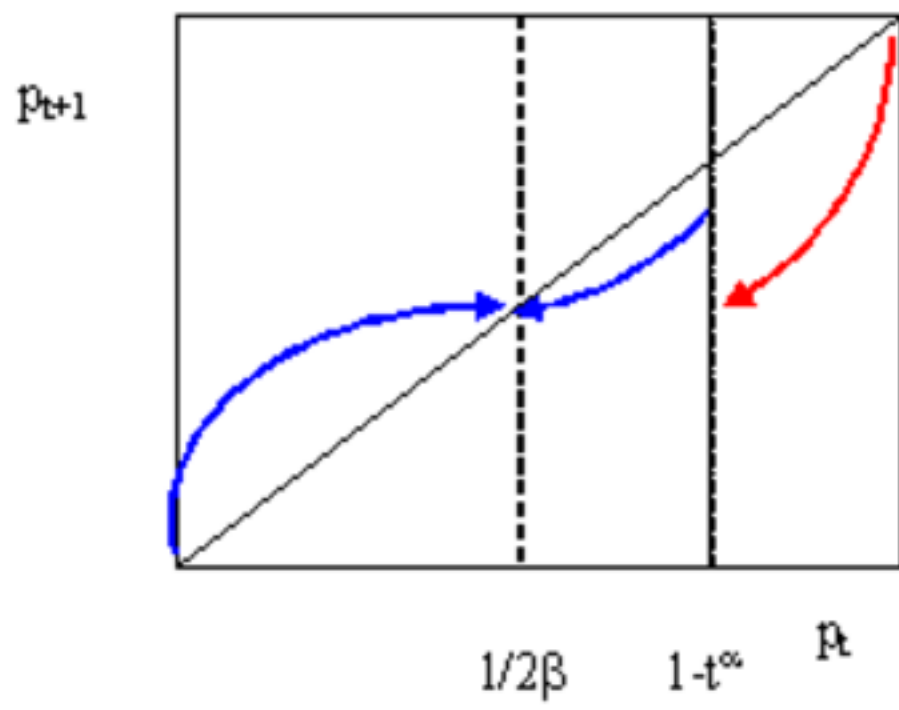


Figure2

